Building The Data Repository For Local-scale Evapotranspiration Data Using Data Science

Yessica Ramirez-Yara*[1], Agnieszka A. Malinowska[1], Ryszard Hejmanowski[1]

[1] AGH University of Krakow, Faculty of Geo-Data Science, Geodesy, and Environmental Engineering, Al. Adama Mickiewicza 30, 30-059 Krakow, Poland

*Corresponding author: yyara@agh.edu.pl

SALGEE & LSA SAF User Workshop on Monitoring Drought Impacts on Vegetation & Feedback
Darmstadt, Germany, 25 - 26 November, 2024

Presentation content

- Introduction
- Materials and Methods
- Results
- Important remarks

Introduction

Climate warming is gradually drying out the land.
Especially in agricultural and forest regions, this leads to a gradual loss of natural and economic potential in many countries.

The rapid surface run-off of rivers after sudden but infrequent rainfall makes it difficult to retain moisture. Evapotranspiration, play a huge role in the drying out of land around the world.

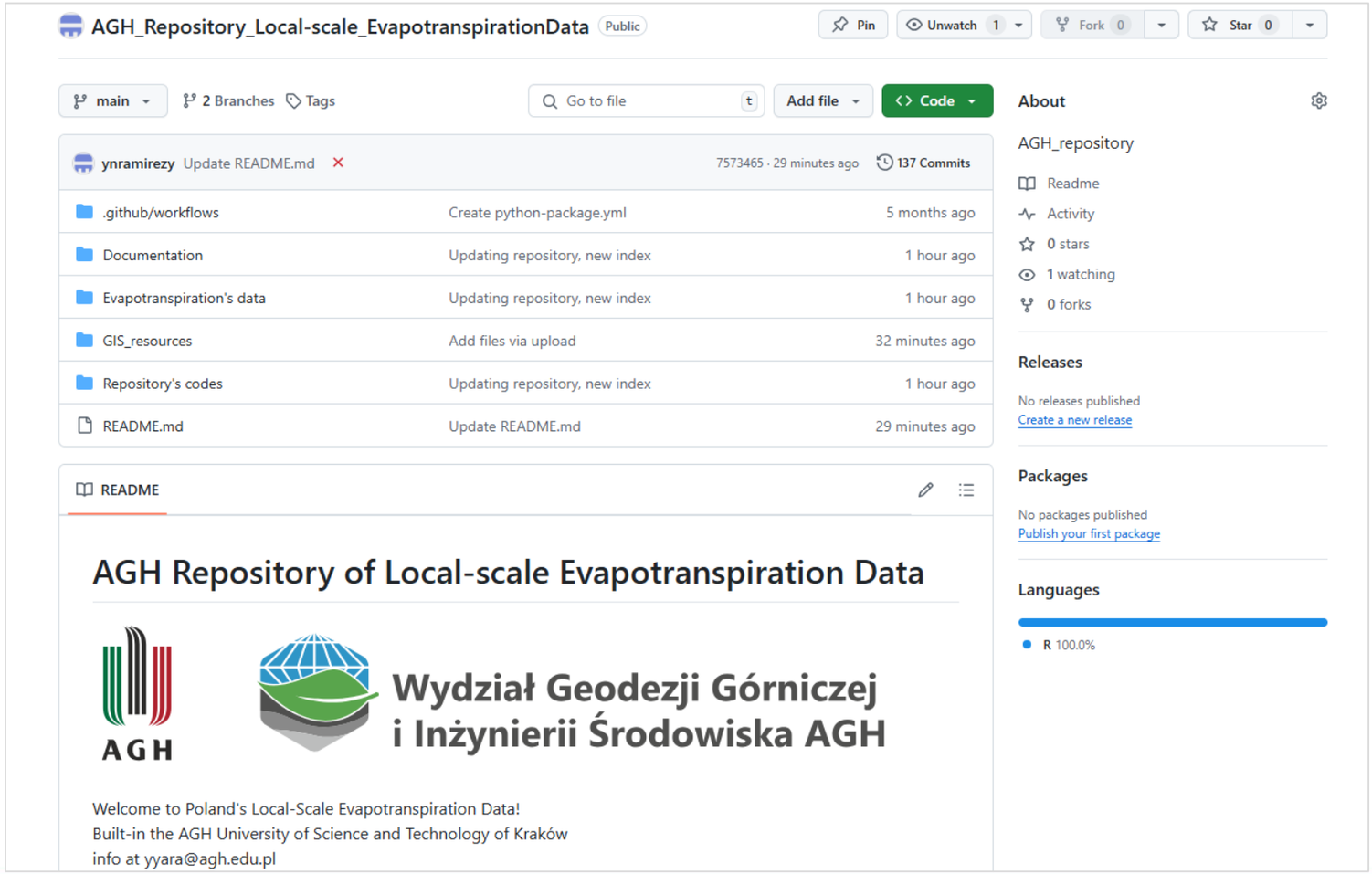
Currently, there are portals that allow visualization of evapotranspiration data through geographic image web services with monthly temporal precision but with restrictions.

Why build a repo?

- Constructing the **CUTTING-EDGE REPOSITORY FOR LOCAL-SCALE EVAPOTRANSPIRATION DATA *IS TO ENSURE THE CENTRALIZATION***, availability and detailed spatial – temporal resolution of integrated, processed, and structured evapotranspiration data along with associated information.
  - The repository development and **DATA INTEGRATION** relied on web scraping, advanced raster geoprocessing tools and spatial statistical analysis thought Anaconda environments.
    - This process was supported by the ***CI/CD*** software concept, utilizing ***Git version*** control tools via ***GitHub***.

¡These evapotranspiration data for Poland, available in the repository, serve as the **first web portal for downloading specialized information for the country**, thus becoming a crucial input for the estimation of runoff and prevention of natural disasters, leveraging the temporal and spatial resolution provided!

1 Repository UI



https://github.com/ynramirezy/AGH_Repository_Local-scale_EvapotranspirationData

Scan to access to the GitHub repository

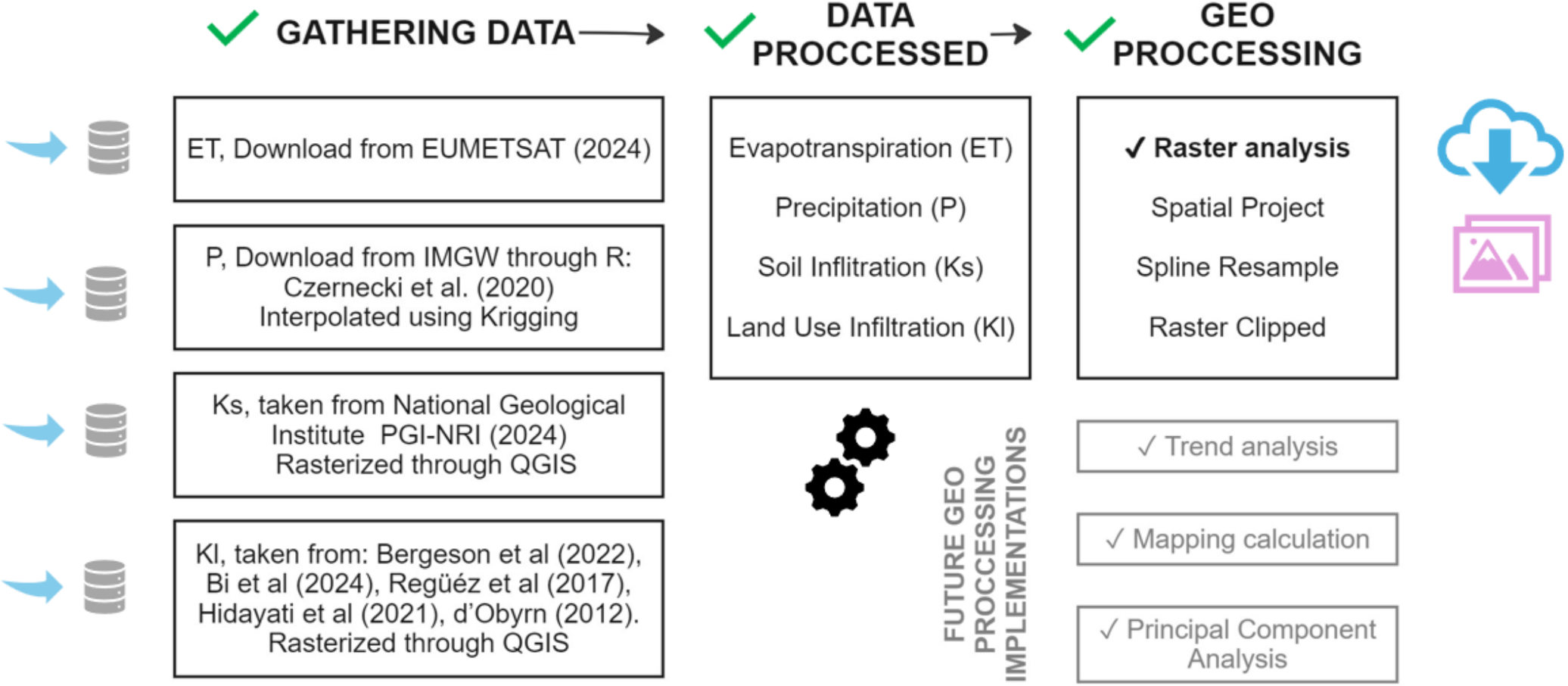
2 Data catalog available in the repository

In the GitHub repository, users can download of data on evapotranspiration, precipitation, soil infiltration, and infiltration by land cover, with the initial conditions outlined. This allows immediate geospatial data analysis using an appropriate matrix format, following the standardization established in the repository.

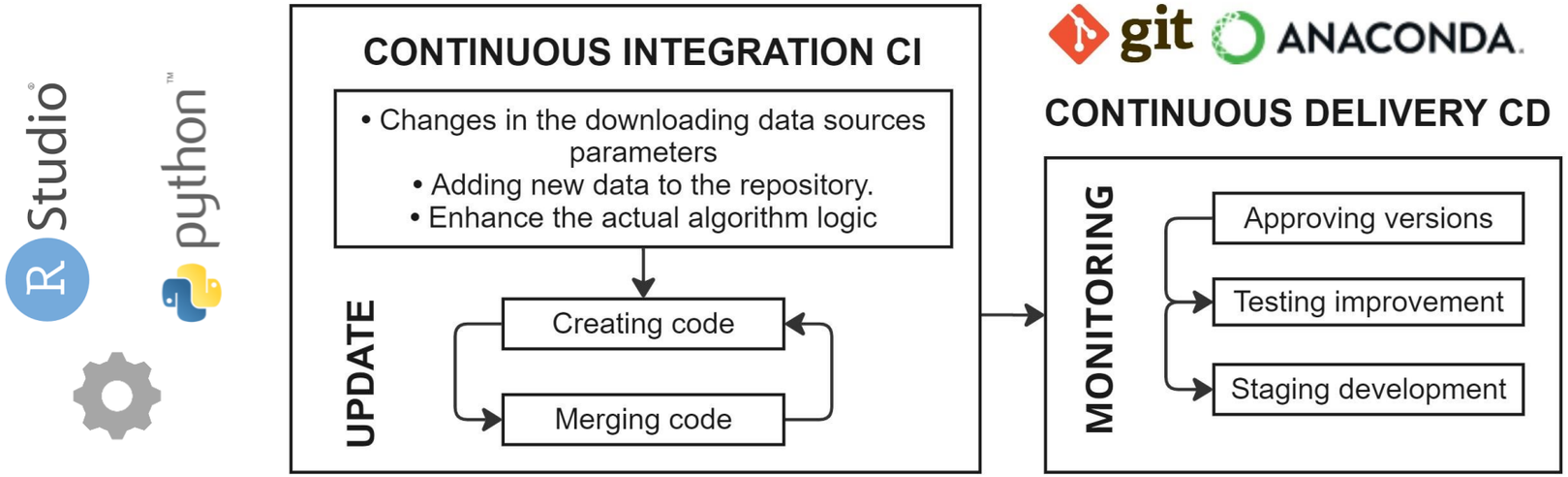
Scan to access to the GitHub repository

Dataset source	Format	Data type	Temporal window	Spatial resolution	Standarization method
Evapotranspiration (EUMETSAT 2024)	.GeoTIFF	Dynamic time indexed	Daily	30-meter pixel (5 counties of Lower Silesian state)	Resampling using BSpline kernel *
Precipitation (Czernecki et al. 2020)	.GeoTIFF	Dynamic time indexed	Daily	30-meter pixel (5 counties of Lower Silesian state)	Interpolation using Ordinary Kriging *
Soil K Infiltration (PGI-NRI 2024)	.GeoTIFF	Static	-	30-meter pixel (Lubin only)	Digitalization of soi infiltration
Land use K Infiltration (Bergeson et al. 2022; Bi et al. 2014; Regüés et al. 2017; Hidayati et al. 2021; d'Obyrn 2012).	.GeoTIFF	Static	-	30-meter pixel (Lubin only)	Land use K infiltratic assignment

3 Repository workflow



4 Repository CI/CD pipelines



- The repository includes code that allows ***users to configure and generate images for various geographic areas***, ranging from municipalities (Gminy) to counties (Powiaty), states (Województwa), and even the entire country.

- The repository is a ***live structure*** that automatically is gathering the data through web scrapping techniques, processing spatial data implying data science advanced tools and updating data ready-for-use in the cloud

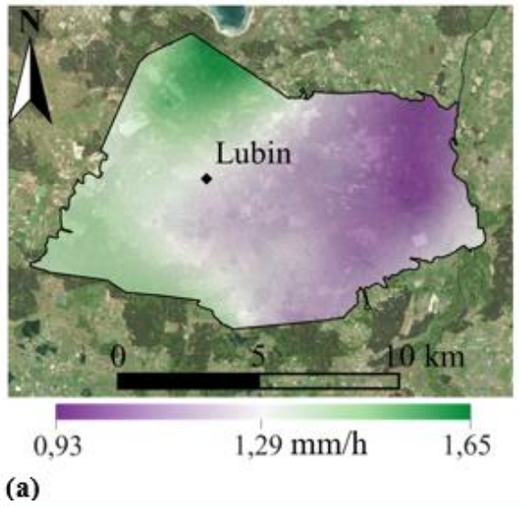
- ***This transforms the repository into not only a centralized hub for evapotranspiration information but also a dynamic portal for generating new data on demand), while the parameter K of soil infiltration, according to land use and soil conditions, is available in the repository only for the city of Lubin.***

5 Sample of repository data of Lubin city for 09/09/2023

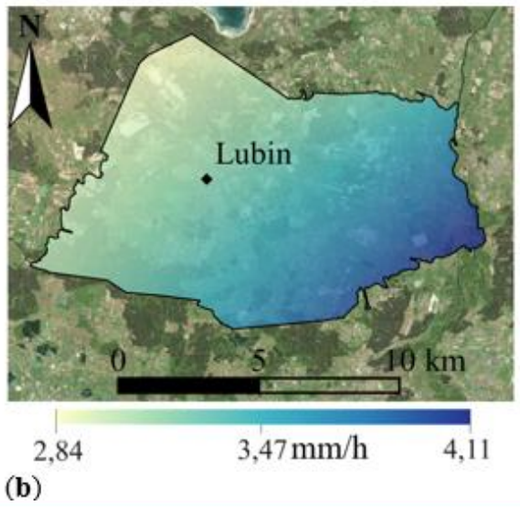
GitHub repository will enable users to directly access geoprocessed data in a common planar reference system, with the same local spatial and temporal resolution for Poland.

The repository contains ready-to-download images for five counties located in the northern part of Lower Silesia (Dolnośląskie): Polkowice, Lubin, Wołów, Legnica, and the city of Legnica.

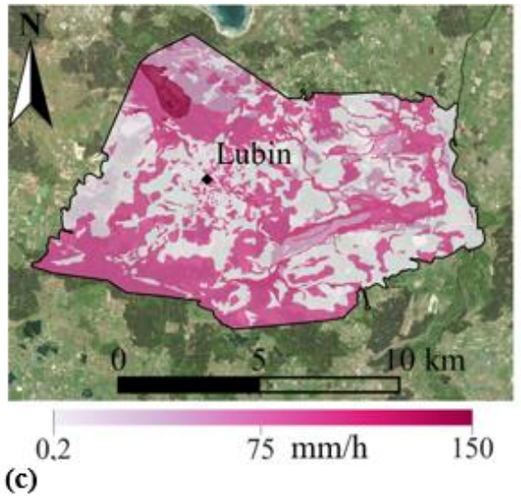
Evapotranspiration



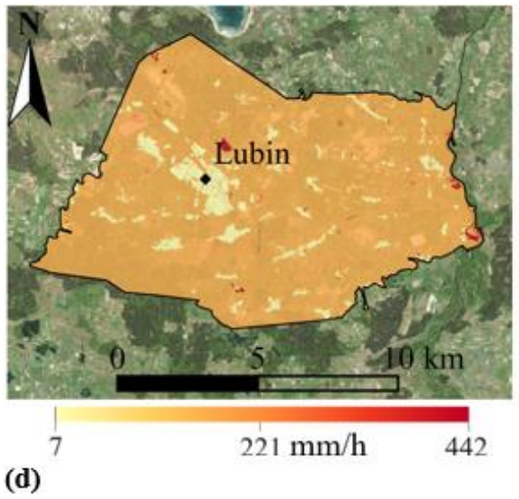
Precipitation



Soil Ks infiltration



Land use Kl



Important remarks, the repository:

- ensures ***LONG-TERM PERSISTENCE AND PRESERVATION OF DATASETS IN THE GEOTIFF FORMAT*** and allows public access and download data and code files, allowing high-quality and high-resolution data to be available measuring the speed of water from when it falls to the ground until it is absorbed by its intrinsic characteristics.

- provides data for download in GeoTIFF raster format, georeferenced to the Polish local high-spatial-temporal reference system, ***ALLOWING IMMEDIATE GEOSPATIAL DATA ANALYSIS IN DIFFERENT PLATFORMS***.

- ***IMPLEMENTS CI/CD PIPELINES IMPROVES DEVELOPMENT TIMES, REDUCES CODE ERRORS IN PRODUCTION, AND OPTIMIZES WORKFLOWS AMONG REPOSITORY COLLABORATORS.***

- becomes a ***HUB FOR CONTINUOUS RESEARCH RELATED TO SOIL EVAPOTRANSPIRATION IN POLAND,*** along with the ongoing updating and enhancement of code, ensuring the availability of data for download.

Building The Data Repository For Local-scale Evapotranspiration Data Using Data Science

Thank you! Vielen Dank! Gracias!

Scan to access to the GitHub repository